Improved Invariant Confidence Intervals for a Normal Variance

Constantinos Goutis[1]

University College London

George Casella[2]

Cornell University

BU-1038-M July, 1989

AMS 1980 *subject classifications.* Primary 62F25; Secondary 62C99.

Key words and phrases. Scale parameter, normal distribution, coverage probability.

[1]This research is part of the author's Ph.D. dissertation at Cornell University.

[2]Research supported by National Science Foundation Grant No. DMS-890039.

ABSTRACT

Confidence intervals for the variance of a normal distribution with unknown mean are constructed which improve upon the usual shortest interval based on the sample variance alone. These intervals have guaranteed coverage probability uniformly greater than a predetermined value $1-\alpha$, and have uniformly shorter length. Using information relating the size of the sample mean to that of the sample variance, we smoothly shift the usual minimum length interval closer to zero, simultaneously bringing the endpoints closer to each other. The construction is generalized to interval estimation of a scale parameter when the location parameter is unknown. We show how to construct the best location-scale invariant interval, and discuss conditions that allow us to improve upon it by requiring only scale invariance. The gains in coverage probability and expected length are also investigated numerically.

1. Introduction. In the problem of estimating the variance of the normal distribution there are two possible cases, depending on whether the mean is known or unknown. When the mean is known the structure of the problem is relatively simple, since by sufficiency the data can be reduced to the sum of squared deviations from the mean and every optimal point or interval estimator must be based on this sufficient statistic. Hodges and Lehmann (1951) proved that the point estimator which is a constant multiple of this sufficient statistic is admissible under squared error loss. For interval estimators Tate and Klett (1959) showed that the endpoints of the shortest $1-\alpha$ confidence interval must be the sum of squared deviations from the mean multiplied by the appropriate constants.

A more complicated problem is that of constructing optimal estimators for the variance of the normal distribution when the mean is unknown. The history of development of solutions can be traced back at least to Stein (1964). Stein showed that we can improve on the "usual" point estimator for the variance by using information in the size of the sample mean relative to the sample variance. His estimation procedure can be thought of as first testing the null hypothesis that the population mean is zero, and, if we accept it, pooling the sample mean and the sample variance. In this way, whenever the population mean seems to be small, we gain another degree of freedom and we are able to beat the estimator based on the sample variance alone for every parameter value. Brown (1968) extended Stein's results to more general loss functions and a larger class of distributions. His estimator is for general scale parameter when the location parameter is unknown. He uses the usual estimator s^2 for scale parameter whenever the estimate of the location parameter y seems large and a smaller multiple of s^2 whenever y seems small. The relative size of s^2 is measured by the statistic $t = y^2/s^2$. Both Stein's and Brown's estimators are inadmissible. Admissible procedures are usually Bayes rules or limits of Bayes rules which must be analytic functions. Therefore it is possible to improve upon these estimators. Brewster and Zidek (1974) were able to find better estimators by taking a finer partition of the set of possible values of t. Their estimator is "smooth" enough to be generalized Bayes and, under some conditions, admissible among scale invariant point estimators. Proskin (1985) showed later that it is admissible within the class of all estimators.

The problem of the interval estimation of variance is in many ways similar to the problem of point estimation. Tate and Klett (1959) calculated the endpoints of the shortest confidence intervals based on s^2 alone. Cohen (1972) was able to construct improved confidence intervals adapting Brown's (1968) techniques. Cohen's intervals keep the same length but, by shifting the endpoints towards zero whenever $t \leq K$, some fixed but arbitrary constant, he was able to dominate Tate and Klett's intervals in terms of coverage probability.

Shorrock (1987) further improved on Cohen's result. In a manner analogous to Brewster and Zidek, Shorrock partitioned the set of possible values of t. By successively adding more cutoff points he was able to construct a "smooth" version of Cohen's interval. The resulting interval is highest posterior density region with respect to an improper prior and dominates the usual interval based on s^2 alone. For both Shorrock and Cohen type intervals the domination is only in terms of coverage probability since, by construction, the length is kept fixed and equal to the usual length. Furthermore, the confidence coefficient remains equal to $1-\alpha$ since asymptotically, as the noncentrality parameter $\lambda = \mu^2/\sigma^2$ tends to infinity, the endpoints of the intervals coincide with the endpoints of the usual interval.

The problem considered in this paper is, in some sense, the dual problem. In Section 2 we construct intervals which improve upon the usual shortest interval based on s^2 alone in terms of length. We keep the minimum coverage probability equal to a predetermined value $1-\alpha$ and we shift the interval closer to zero whenever the sample indicates that the mean is close to zero. By shifting we are able to bring the endpoints closer to each other, hence producing shorter intervals. Using a method similar to Brewster and Zidek we construct a family of "smooth" $(1-\alpha)\,100$ % intervals which are shorter than the usual interval and, consequently, Cohen and Shorrock type intervals. In Section 3, we first generalize conditions for constructing optimal location-scale invariant intervals, making as few distributional assumptions as possible. Then we show that, under suitable assumptions, we can use techniques similar to Section 2 to construct families of intervals for the scale parameter when the location parameter is unknown.

2. Normal variance. Let $\underset{\sim}{X} = (\underset{\sim}{X}_1, \underset{\sim}{X}_2)$ be a $(n+p)\times 1$ vector so that $\underset{\sim}{X}_1 = (X_1, X_2, \ldots, X_n)$ and $\underset{\sim}{X}_2 = (X_{n+1}, \ldots, X_{n+p})$. We assume that $\underset{\sim}{X}$ is a random variable from a multivariate normal distribution with mean $(\underset{\sim}{0}, \underset{\sim}{\mu})$ where $\underset{\sim}{0}$ is a vector of order n and $\underset{\sim}{\mu} = (\mu_1, \mu_2, \ldots, \mu_p)$ is unknown, and covariance matrix σ^2 times the identity matrix of order $n+p$. We are interested in estimating the unknown parameter σ^2.

Let $s^2 = \underset{\sim}{X}_1{}' \underset{\sim}{X}_1$, $y^2 = \underset{\sim}{X}_2{}' \underset{\sim}{X}_2$ and $t = y^2/s^2$. By sufficiency the data can be reduced to $(s^2, \underset{\sim}{X}_2)$. With the normality assumption we have that

$$\frac{s^2}{\sigma^2} \sim \chi^2_n \tag{2.1}$$

and

$$\frac{y^2}{\sigma^2} \sim \chi^2_p(\lambda), \quad \lambda = \frac{\underset{\sim}{\mu}'\underset{\sim}{\mu}}{\sigma^2}, \tag{2.2}$$

a central and noncentral chi squared distribution, the latter with noncentrality parameter λ. The noncentral chi squared density with n degrees of freedom and noncentrality parameter λ will be denoted by $f_n(x\,;\,\lambda)$. If $\lambda = 0$ we will omit λ from the notation and $f_n(x)$ will be the central chi squared density. The respective cumulative distribution functions will be denoted by $F_n(x\,;\,\lambda)$ and $F_n(x)$.

We can think of the setup of the problem as the general linear hypothesis, where y^2 represents the model sum of squares and s^2 the error sum of squares in an analysis of variance table. For the simple setup of estimating a variance from a sample $X_1, X_2, \ldots, X_N$ from a single normal population with unknown mean, $s^2 = \sum(X_i - \overline{X})^2$ and $y^2 = N\overline{X}^2$ where

$\overline{X} = \sum X_i / N$. Then we have $N-1$ central and 1 noncentral degrees of freedom and the noncentrality parameter is equal to μ^2/σ^2.

The minimum length intervals, based on s^2 alone were tabulated by Tate and Klett (1959) and have the form

$$C_U(s^2) = \left(\frac{1}{b_n}s^2, \frac{1}{a_n}s^2\right) \tag{2.3}$$

where a_n and b_n satisfy

$$\int_{a_n}^{b_n} f_n(x)\,dx = 1-\alpha \tag{2.4}$$

and

$$f_{n+4}(a_n) = f_{n+4}(b_n). \tag{2.5}$$

Let K be a positive constant and $\tau(x)$ be an increasing continuous function defined on $(0, +\infty)$ such that $\tau(x) > x$ for every x. Define a confidence procedure as follows:

$$I_1(s^2, t, K) = \begin{cases} \left(\frac{1}{b_n}s^2, \frac{1}{a_n}s^2\right) & \text{if} \quad t > K \\ \\ (\phi_1(K)s^2, \phi_2(K)s^2) & \text{if} \quad t \leq K \end{cases}$$

where $\phi_1(K)$ and $\phi_2(K)$ are determined from the following equations

$$\int_{\phi_1(K)}^{\phi_2(K)} f_{n+4}\left(\tfrac{1}{x}\right) F_p\left(\tfrac{K}{x}\right) dx = \int_{\frac{1}{b_n}}^{\frac{1}{a_n}} f_{n+4}\left(\tfrac{1}{x}\right) F_p\left(\tfrac{K}{x}\right) dx \tag{2.6}$$

and

$$f_{n+4}\left\{\frac{1}{\phi_1(K)}\right\} F_p\left\{\frac{\tau(K)}{\phi_1(K)}\right\} = f_{n+4}\left\{\frac{1}{\phi_2(K)}\right\} F_p\left\{\frac{\tau(K)}{\phi_2(K)}\right\}. \tag{2.7}$$

If $K = +\infty$ the interval coincides with the usual one since $\tau(+\infty) = +\infty$. Note that the procedure defines a class of confidence intervals rather than a single interval since the endpoints depend on an unspecified function τ. For any K and a given functional form of τ we can choose $\phi_1(K)$ and $\phi_2(K)$ in a unique way. For the most part, the values of K are only a means to an end, and for clarity of notation we will sometimes omit K from the notation whenever no confusion arises.

The way of constructing $I_1(s^2, t, K)$ is analogous to Brown's (1968) point estimator and Cohen's (1972) confidence interval. We partition the space of possible values of t, and whenever t is smaller than a constant we shift the endpoints towards zero. In our case we keep the coverage probability, under $\underset{\sim}{\mu} = \underset{\sim}{0}$, equal to $1-\alpha$. The following theorems establish that the new interval improves upon $C_U(s^2)$.

THEOREM 2.1. The coverage probability of the procedure $I_1(s^2, t, K)$ is greater than the coverage probability of $C_U(s^2)$. The probability is strictly greater if $\lambda > 0$.

PROOF. Note that the intervals differ only when $t \le K$. Working with the joint probability it suffices to show that

$$P\{\sigma^2 \in (\phi_1 s^2, \phi_2 s^2), t \le K\} \ge P\{\sigma^2 \in (\tfrac{1}{b_n} s^2, \tfrac{1}{a_n} s^2), t \le K\}. \tag{2.8}$$

Now observe that y^2 and s^2 are independent and y^2/σ^2 has a noncentral chi squared distribution with p degrees of freedom and noncentrality parameter λ, and s^2/σ^2 has a chi squared distribution with n degrees of freedom. Using the conditional density of $\sigma^2/s^2 = x$, inequality (2.8) can be written as

$$\int_{\phi_1}^{\phi_2} f_{n+4}(\tfrac{1}{x})\, F_p(\tfrac{K}{x}; \lambda)\, dx \ge \int_{\frac{1}{b_n}}^{\frac{1}{a_n}} f_{n+4}(\tfrac{1}{x})\, F_p(\tfrac{K}{x}; \lambda)\, dx. \tag{2.9}$$

For $\lambda = 0$, (2.9) is an equality. For $\lambda > 0$ we will show that we have strict inequality.

For fixed γ, λ define the function $g_{\gamma,\lambda}(w)$ as the solution to

$$\gamma = \int_{w}^{g_{\gamma,\lambda}(w)} f_{n+4}(\tfrac{1}{x})\, F_p(\tfrac{K}{x}; \lambda)\, dx. \tag{2.10}$$

Note that $g_{\gamma,\lambda}(w) \ge w$ and $g_{\gamma,\lambda}(w)$ is increasing in w. Furthermore let

$$\gamma_1 = \int_{\frac{1}{b_n}}^{\frac{1}{a_n}} f_{n+4}(\tfrac{1}{x})\, F_p(\tfrac{K}{x})\, dx \tag{2.11}$$

and

$$\gamma_2 = \int_{\frac{1}{b_n}}^{\frac{1}{a_n}} f_{n+4}(\tfrac{1}{x})\, F_p(\tfrac{K}{x}\,;\lambda)\, dx. \tag{2.12}$$

For fixed γ and λ, since $d\gamma/dw = 0$ we have

$$\frac{dg(w)}{dw} f_{n+4}(\frac{1}{g(w)})\, F_p(\frac{K}{g(w)};\lambda) - f_{n+4}(\tfrac{1}{w})\, F_p(\tfrac{K}{w};\lambda) = 0, \tag{2.13}$$

which implies

$$\frac{dg(w)}{dw} = \frac{f_{n+4}(\frac{1}{w})\, F_p(\frac{K}{w};\lambda)}{f_{n+4}(\frac{1}{g(w)})\, F_p(\frac{K}{g(w)};\lambda)}. \tag{2.14}$$

We can establish (2.9) by showing $\phi_2 > g_{\gamma_2,\lambda}(\phi_1)$.

Define $G(w) = g_{\gamma_1,0}(w) - g_{\gamma_2,\lambda}(w)$ and note that $G(\phi_1) = \phi_2 - g_{\gamma_2,\lambda}(\phi_1)$. We will prove that $\phi_2 - g_{\gamma_2,\lambda}(\phi_1)$ is positive by showing that G satisfies the assumptions of Lemma A.1, that is, if G' is negative when G is zero, then G has only one sign change from positive to negative.

First of all we see that $g_{\gamma_1,0}(1/b_n) = g_{\gamma_2,\lambda}(1/b_n) = 1/a_n$ hence $G(1/b_n)$ is equal to zero. Let x_0 be a point such that $G(x_0)=0$ and let $y_0 = g_{\gamma_1,0}(x_0) = g_{\gamma_2,\lambda}(x_0)$. Then using (2.14), we get after some simplification

$$\left.\frac{dG(w)}{dw}\right|_{w=x_0} = \frac{f_{n+4}(\frac{1}{x_0})}{f_{n+4}(\frac{1}{y_0})} \left\{ \frac{F_p(\frac{K}{x_0})}{F_p(\frac{K}{y_0})} - \frac{F_p(\frac{K}{x_0};\lambda)}{F_p(\frac{K}{y_0};\lambda)} \right\}. \tag{2.15}$$

Since the chi squared distribution has monotone likelihood ratio in the noncentrality parameter we conclude that $F_p(x\,;\lambda)/F_p(x)$ is increasing in x. Since $x_0 < y_0$, the term in braces in (2.15) is less than zero, thus the derivative evaluated at x_0 is negative. Also, $G(1/b_n) = 0$ so we have that G is positive for every number less than $1/b_n$. From Lemma A.2, $\phi_1 < 1/b_n$ therefore $\phi_2 > g_{\gamma_2,\lambda}(\phi_1)$ and (2.12) is established. □

THEOREM 2.2. The length of the interval $I_1(s^2, t, K)$ is, with positive probability, smaller than the length of the usual minimum length interval $C_U(s^2)$.

PROOF. When $t \le K$, the length of the confidence interval, $(\phi_2 - \phi_1)s^2$, is equal to $(g_{\gamma_1,0}(w) - w)s^2$ by (2.6) and (2.11). Unimodality of $f_{n+4}(1/x)\, F_p(K/x)$ (Lemma A.4) implies that the length, as a function of w, has a unique minimum at ϕ_1^0 where ϕ_1^0 and ϕ_2^0 denote the numbers that satisfy

$$\int_{\phi_1^0}^{\phi_2^0} f_{n+4}(\tfrac{1}{x})\, F_p(\tfrac{K}{x})\, dx = \int_{\frac{1}{b_n}}^{\frac{1}{a_n}} f_{n+4}(\tfrac{1}{x})\, F_p(\tfrac{K}{x})\, dx \tag{2.16}$$

and

$$f_{n+4}(\frac{1}{\phi_1^0})\, F_p(\frac{K}{\phi_1^0}) = f_{n+4}(\frac{1}{\phi_2^0})\, F_p(\frac{K}{\phi_2^0}). \tag{2.17}$$

since at the point ϕ_1^0 we have

$$\frac{d\left[g_{\gamma_1,0}(w) - w\right]}{dw}\bigg|_{w=\phi_1^0} = \frac{f_{n+4}(\frac{1}{\phi_1^0})\, F_p(\frac{K}{\phi_1^0})}{f_{n+4}(\frac{1}{\phi_2^0})\, F_p(\frac{K}{\phi_2^0})} - 1 = 0. \tag{2.18}$$

In order to prove that $\phi_2 - \phi_1$ is smaller than $(1/a_n) - (1/b_n)$ it would suffice to show that $\phi_1^0 < \phi_1 < 1/b_n$. From Lemma A.2 $\phi_1 < 1/b_n$, so the result follows if the derivative of length evaluated at ϕ_1 is positive. This is equivalent to

$$f_{n+4}(\frac{1}{\phi_1})\, F_p(\frac{K}{\phi_1}) > f_{n+4}(\frac{1}{\phi_2})\, F_p(\frac{K}{\phi_2}). \tag{2.19}$$

Because of (2.7) the last inequality holds if and only if

$$\frac{F_p\left\{\frac{K}{\phi_1}\right\}}{F_p\left\{\frac{\tau(K)}{\phi_1}\right\}} > \frac{F_p\left\{\frac{K}{\phi_2}\right\}}{F_p\left\{\frac{\tau(K)}{\phi_2}\right\}} \tag{2.20}$$

which can be seen to be true by applying Lemma A.5 with $x_1 = \tau(K)/\phi_1$, $x_2 = \tau(K)/\phi_2$ and $\beta = K/\tau(K)$. Note that β is smaller than 1 because we have assumed $\tau(K) > K$, and Lemma A.5 exploits the fact that the chi square densities have monotone likelihood ratio increasing in the scale parameter. $\square$

The coverage probability of the procedure $I_1(s^2, t, K)$ depends on the unknown parameter λ and the length is actually a random variable depending on s^2 and t. Neither the length nor the coverage probability depends on σ^2. The procedure can be further improved if our action depends on two cutoff points instead of one. The technique of partitioning the set of possible values of t has been implemented in the construction of a point estimator for the variance by Brewster and Zidek (1974) and in the confidence interval constructed by Shorrock (1987).

Given $\underset{\sim}{K}_2 = (K_1, K_2)$, $K_2 < K_1$, define the confidence procedure

$$I_2(s^2, t, \underset{\sim}{K}_2) = \begin{cases} \left(\frac{1}{b_n}s^2, \frac{1}{a_n}s^2\right) & \text{if} \quad t > K_1 \\ (\phi_1(K_1)s^2, \phi_2(K_1)s^2) & \text{if} \quad K_2 < t \le K_1 \\ (\phi_1(K_2)s^2, \phi_2(K_2)s^2) & \text{if} \quad t \le K_2, \end{cases}$$

where ϕ_1 and ϕ_2 satisfy the equations:

$$\int_{\phi_1(K_1)}^{\phi_2(K_1)} f_{n+4}\left(\tfrac{1}{x}\right) F_p\left(\tfrac{K_1}{x}\right) dx = \int_{\frac{1}{b_n}}^{\frac{1}{a_n}} f_{n+4}\left(\tfrac{1}{x}\right) F_p\left(\tfrac{K_1}{x}\right) dx \tag{2.21}$$

$$\int_{\phi_1(K_2)}^{\phi_2(K_2)} f_{n+4}\left(\tfrac{1}{x}\right) F_p\left(\tfrac{K_2}{x}\right) dx = \int_{\phi_1(K_1)}^{\phi_2(K_1)} f_{n+4}\left(\tfrac{1}{x}\right) F_p\left(\tfrac{K_2}{x}\right) dx \tag{2.22}$$

$$f_{n+4}\left\{\frac{1}{\phi_1(K_i)}\right\} F_p\left\{\frac{\tau(K_i)}{\phi_1(K_i)}\right\} = f_{n+4}\left\{\frac{1}{\phi_2(K_i)}\right\} F_p\left\{\frac{\tau(K_i)}{\phi_2(K_i)}\right\} \tag{2.23}$$

for $i = 1, 2$.

THEOREM 2.3. The coverage probability of $I_2\ (s^2, t, \underset{\sim}{K}_2)$ is greater than the coverage probability of $I_1\ (s^2, t, K_1)$.

PROOF. Observing that the two intervals $I_1\ (s^2, t, K_1)$ and $I_2\ (s^2, t, \underset{\sim}{K}_2)$ differ only if $t \leq K_2$, it suffices to show

$$\int_{\phi_1(K_2)}^{\phi_2(K_2)} f_{n+4}(\tfrac{1}{x})\ F_p(\tfrac{K_2}{x};\lambda)\ dx \geq \int_{\phi_1(K_1)}^{\phi_2(K_1)} f_{n+4}(\tfrac{1}{x})\ F_p(\tfrac{K_2}{x};\lambda)\ dx. \tag{2.24}$$

By Lemma A.3 we know that the interval is further shifted towards zero, that is, $\phi_1(K_1) > \phi_1(K_2)$. The result follows as in Theorem 2.1 by replacing K, $1/b_n$ and $1/a_n$ by K_2, $\phi_1(K_1)$ and $\phi_2(K_1)$ respectively. □

THEOREM 2.4. If $t \leq K_2$ the length of $I_2\ (s^2, t, \underset{\sim}{K}_2)$ is smaller than the length of $I_1\ (s^2, t, K_1)$.

PROOF. As in Theorem 2.2 the length, as a function of the lower limit of the integration subject to (2.22), has a unique minimum at $\phi_1^0(K_2)$ where $\phi_1^0(K_2)$ and $\phi_2^0(K_2)$ are defined in way analogous to (2.16) and (2.17), that is,

$$f_{n+4}\left\{\frac{1}{\phi_1^0(K_2)}\right\} F_p\left\{\frac{K_2}{\phi_1^0(K_2)}\right\} = f_{n+4}\left\{\frac{1}{\phi_2^0(K_2)}\right\} F_p\left\{\frac{K_2}{\phi_2^0(K_2)}\right\} \tag{2.25}$$

and

$$\int_{\phi_1^0(K_2)}^{\phi_2^0(K_2)} f_{n+4}(\tfrac{1}{x})\ F_p(\tfrac{K_2}{x})\ dx = \int_{\phi_1(K_1)}^{\phi_2(K_1)} f_{n+4}(\tfrac{1}{x})\ F_p(\tfrac{K_2}{x})\ dx. \tag{2.26}$$

In order to have $\phi_1(K_1) > \phi_1(K_2) > \phi_1^0(K_2)$, it suffices that

$$\left.\frac{d\left[g_{\gamma,0}(l) - w\right]}{dw}\right|_{w=\phi_1(K_2)} > 0 \tag{2.27}$$

where

$$\gamma = \int_{w}^{g_{\gamma,0}(w)} f_{n+4}(\tfrac{1}{x})\ F_p(\tfrac{K_2}{x})\ dx. \tag{2.28}$$

Using (2.23) with $i = 2$, the derivative of $g_{\gamma,0}(w)$ is positive if and only if

$$\frac{F_p\left\{ \dfrac{K_2}{\phi_1(K_2)} \right\}}{F_p\left\{ \dfrac{\tau(K_2)}{\phi_1(K_2)} \right\}} > \frac{F_p\left\{ \dfrac{K_2}{\phi_2(K_2)} \right\}}{F_p\left\{ \dfrac{\tau(K_2)}{\phi_2(K_2)} \right\}}. \tag{2.29}$$

Now we can apply Lemma A.5 with $x_1 = \tau(K_2)/\phi_1(K_2)$, $x_2 = \tau(K_2)/\phi_2(K_2)$ and $\beta = K_2/\tau(K_2)$ which is smaller than 1 since $\tau(K) > K$ by assumption, and conclude that equation (2.29) is true, proving the theorem. □

We can easily generalize and improve on $\mathbf{I_2}$ $(s^2, t, \underset{\sim}{K}_2)$ by taking three cutoff points etc. In general, any interval based on a finite number cutoff points can be improved by adding an extra cutoff point. Working as in Brewster and Zidek (1974) and Shorrock (1987) we can create a triangular array $\{\underset{\sim}{K}_m\}$ that will fill up the interval $(0, +\infty)$ and take the confidence interval that is the limit of the confidence intervals based on $\underset{\sim}{K}_m$, as m tends to $+\infty$. It is plausible that the limiting interval will be better in terms of length than the usual minimum length interval. However, the form of the limiting interval is not obvious, since as we can see in equations (2.21)–(2.23) the numbers $\phi_1(K_2)$ and $\phi_2(K_2)$ depend not only on K_2 and the function τ but also on K_1. Hence for a given t the endpoints of the interval depend on all the cutoff points K_i that are greater than or equal to t.

We create a triangular array $\{\underset{\sim}{K}_m\}$ array as follows: For each m, define $\underset{\sim}{K}_m = (K_{m,1}, \ldots, K_{m,m-1}, K_{m,m})$, where $0 < K_{m,1} < \ldots < K_{m,m-1} < K_{m,m} < +\infty$. Furthermore, we require $\lim_{m\to\infty} K_{m,1} = 0$ and $\lim_{m\to\infty} K_{m,m} = +\infty$ and $\lim_{m\to\infty} \max_i (K_{m,i} - K_{m,i-1}) = 0$

As $m \to \infty$ the endpoints of the intervals based on $\underset{\sim}{K}_m$ tend to some functions $\phi_1(t)$ and $\phi_2(t)$. In order to determine $\phi_1(t)$ and $\phi_2(t)$ we define

$$K_{m,i(t)} = \inf\{ K \in \underset{\sim}{K}_m : K \geq t \}. \tag{2.30}$$

Then for given t and s the confidence interval at the mth stage is

$(\phi_1(K_{m,i(t)})s^2, \phi_2(K_{m,i(t)})s^2)$ where $\phi_1(K_{m,i(t)})$ and $\phi_2(K_{m,i(t)})$ satisfy

$$\int_{\phi_1(K_{m,i(t)})}^{\phi_2(K_{m,i(t)})} f_{n+4}\left(\tfrac{1}{x}\right) F_p\left(\frac{K_{m,i(t)}}{x}\right) dx = \int_{\phi_1(K_{m,i(t)+1})}^{\phi_2(K_{m,i(t)+1})} f_{n+4}\left(\tfrac{1}{x}\right) F_p\left(\frac{K_{m,i(t)}}{x}\right) dx \tag{2.31}$$

and

$$f_{n+4}\left\{\frac{1}{\phi_1(K_{m,i(t)})}\right\} F_p\left\{\frac{\tau(K_{m,i(t)})}{\phi_1(K_{m,i(t)})}\right\} = f_{n+4}\left\{\frac{1}{\phi_2(K_{m,i(t)})}\right\} F_p\left\{\frac{\tau(K_{m,i(t)})}{\phi_2(K_{m,i(t)})}\right\}. \tag{2.32}$$

Now use Taylor's theorem and substitute $F_p(K_{m,i(t)}/x)$ in the RHS of (2.31) by its series expansion around $K_{m,i(t)+1}$, keeping the first two terms. We get

$$\int_{\phi_1(K_{m,i(t)})}^{\phi_2(K_{m,i(t)})} f_{n+4}\left(\tfrac{1}{x}\right) F_p\left(\frac{K_{m,i(t)}}{x}\right) dx \tag{2.33}$$

$$= \int_{\phi_1(K_{m,i(t)+1})}^{\phi_2(K_{m,i(t)+1})} f_{n+4}\left(\tfrac{1}{x}\right) \left\{ F_p\left(\frac{K_{m,i(t)+1}}{x}\right) + (K_{m,i(t)} - K_{m,i(t)+1}) \times \right.$$

$$\left. \tfrac{1}{x} f_p\left(\frac{K_{m,i(t)+1}}{x}\right) + \frac{\left(K_{m,i(t)} - K_{m,i(t)+1}\right)^2}{2} \left(\tfrac{1}{x}\right)^2 f_p{}'\left(\frac{K_r}{x}\right) \right\} dx$$

where K_r is some number in the interval $(K_{m,i(t)}, K_{m,i(t)+1})$. Bring the first term of the sum of the RHS to the left of the equality, divide both sides by $K_{m,i(t)} - K_{m,i(t)+1}$ and take the limit as m goes to $+\infty$. By the construction of the array the difference $K_{m,i(t)} - K_{m,i(t)+1}$ tends to zero and $K_{m,i(t)}$ tends to t. Hence equation (2.33) becomes, as $m \to +\infty$,

$$\frac{d}{dt}\left[\int_{\phi_1(t)}^{\phi_2(t)} f_{n+4}\left(\frac{1}{x}\right) F_p\left(\frac{t}{x}\right) dx\right] = \int_{\phi_1(t)}^{\phi_2(t)} f_{n+4}\left(\frac{1}{x}\right) \frac{1}{x} f_p\left(\frac{t}{x}\right) dx. \tag{2.34}$$

The limit of the LHS is justified because, in the limit, the remainder term disappears. Since the derivative $f_p{}'$ is bounded in finite intervals, the integral

$$\int_{\phi_1(K_{m,i(t)+1})}^{\phi_2(K_{m,i(t)+1})} f_{n+4}\left(\frac{1}{x}\right)\left(\frac{1}{x}\right)^2 f_p{}'\left(\frac{K_r}{x}\right) dx \tag{2.35}$$

is also bounded. Hence $\lim_{m\to\infty}\left(K_{m,i(t)} - K_{m,i(t)+1}\right) = 0$ implies that the remainder tends to zero. Using Leibniz' formula for the differentiation of the integral the last relation becomes

$$\frac{d\phi_1(t)}{dt} f_{n+4}\left\{\frac{1}{\phi_1(t)}\right\} F_p\left\{\frac{t}{\phi_1(t)}\right\} = \frac{d\phi_2(t)}{dt} f_{n+4}\left\{\frac{1}{\phi_2(t)}\right\} F_p\left\{\frac{t}{\phi_2(t)}\right\}. \tag{2.36}$$

On the other hand, since we have assumed that the function τ is continuous, relation (2.32) becomes

$$f_{n+4}\left\{\frac{1}{\phi_1(t)}\right\} F_p\left\{\frac{\tau(t)}{\phi_1(t)}\right\} = f_{n+4}\left\{\frac{1}{\phi_2(t)}\right\} F_p\left\{\frac{\tau(t)}{\phi_2(t)}\right\}. \tag{2.37}$$

In order to solve equations (2.36) and (2.37) for ϕ_1 and ϕ_2 we need initial conditions which are given by the equalities

$$\lim_{t\to\infty} \phi_1(t) = \frac{1}{b_n} \text{ and } \lim_{t\to\infty} \phi_2(t) = \frac{1}{a_n}. \tag{2.38}$$

It is obvious that for different forms of the function τ we have different confidence intervals. By the Lebesgue Dominated Convergence Theorem we know that the confidence coefficient of any interval constructed in this way is $1-\alpha$. For the length of the limiting intervals we have the following result.

THEOREM 2.5. For every $t < \infty$ the length $(\phi_2(t) - \phi_1(t))s^2$ is smaller than $((1/a_n) - (1/b_n))s^2$.

PROOF. Rearranging equations (2.36) and (2.37) yields

$$\frac{d\,\phi_2(t)}{dt} \frac{F_p\left\{\frac{t}{\phi_2(t)}\right\}}{F_p\left\{\frac{\tau(t)}{\phi_2(t)}\right\}} = \frac{d\,\phi_1(t)}{dt} \frac{F_p\left\{\frac{t}{\phi_1(t)}\right\}}{F_p\left\{\frac{\tau(t)}{\phi_1(t)}\right\}}. \tag{2.39}$$

Using Lemma A.5 with $x_1 = \tau(t)/\phi_1(t)$, $x_2 = \tau(t)/\phi_2(t)$, $\beta = t/\tau(t)$ we have

$$\frac{F_p\left\{\frac{t}{\phi_2(t)}\right\}}{F_p\left\{\frac{\tau(t)}{\phi_2(t)}\right\}} < \frac{F_p\left\{\frac{t}{\phi_1(t)}\right\}}{F_p\left\{\frac{\tau(t)}{\phi_1(t)}\right\}} \tag{2.40}$$

therefore (2.39) implies that $d\left[\phi_2(t) - \phi_1(t)\right]/dt > 0$, that is, the length is an increasing function of t. But we know that

$$\lim_{t\to\infty} [\, \phi_2(t) - \phi_1(t) \,] = \frac{1}{a_n} - \frac{1}{b_n} \tag{2.41}$$

so for any $t < +\infty$ the length is strictly smaller than $((1/a_n) - (1/b_n))s^2$. □

It is interesting to see how the endpoints of the intervals degenerate in some special forms of the function $\tau(K)$. If $\tau(K)$ is equal to infinity then equation (2.7) becomes

$$f_{n+4}\left\{\frac{1}{\phi_1(K)}\right\} = f_{n+4}\left\{\frac{1}{\phi_2(K)}\right\} \tag{2.42}$$

and, together with equation (2.6), implies that $\phi_1(K)$ and $\phi_2(K)$ coincide with $1/b_n$ and $1/a_n$ respectively. Therefore the interval based on one cutoff point is identical to $C_U(s^2)$. By taking more cutoff points we do not shift the endpoints, therefore the limiting interval coincides with the usual minimum length confidence interval based on s^2 alone.

On the other hand if we take $\tau(K) = K$ the endpoints are $\phi_1^0(K)$ and $\phi_2^0(K)$ defined by equations (2.16) and (2.17). It is tempting to choose $\tau(K) = K$, since, if we do so, we maximize the gain in terms of length. However, by filling $(0, +\infty)$ with cutoff points, the endpoints of the limiting interval are defined by the equations

$$\frac{d\phi_1(t)}{dt}\, f_{n+4}\left\{\frac{1}{\phi_1(t)}\right\} F_p\left\{\frac{t}{\phi_1(t)}\right\} = \frac{d\phi_2(t)}{dt}\, f_{n+4}\left\{\frac{1}{\phi_2(t)}\right\} F_p\left\{\frac{t}{\phi_2(t)}\right\} \tag{2.43}$$

and

$$f_{n+4}\left\{\frac{1}{\phi_1(t)}\right\} F_p\left\{\frac{t}{\phi_1(t)}\right\} = f_{n+4}\left\{\frac{1}{\phi_2(t)}\right\} F_p\left\{\frac{t}{\phi_2(t)}\right\} \tag{2.44}$$

and the initial conditions (2.38). Equations (2.43) and (2.44) imply

$$\frac{d\,\phi_2(t)}{dt} = \frac{d\,\phi_1(t)}{dt}, \tag{2.45}$$

hence the interval has constant length. Because of the initial conditions, we can conclude that

$$\phi_2(t) - \phi_1(t) = \frac{1}{a_n} - \frac{1}{b_n} = c_0. \tag{2.46}$$

Substituting $\phi_1(t) + c_0$ for $\phi_2(t)$ equation (2.44) becomes

$$f_{n+4}\left\{\frac{1}{\phi_1(t)}\right\} F_p\left\{\frac{t}{\phi_1(t)}\right\} = f_{n+4}\left\{\frac{1}{\phi_1(t)+c_0}\right\} F_p\left\{\frac{t}{\phi_1(t)+c_0}\right\} \tag{2.47}$$

which, surprisingly enough, is the defining equation for Shorrock's interval. It is interesting to note that every interval, based on a finite number on cutoff points has length less than the usual or Shorrock's interval but the limiting length is equal to $c_0\, s^2$.

The same procedure can be implemented to construct improved confidence intervals using the ratio of endpoints as a measure of volume. The minimum ratio intervals based on s^2 alone, tabulated by Tate and Klett (1959), have endpoints satisfying

$$\int_{a_n}^{b_n} f_n(x)\, dx = 1-\alpha \tag{2.48}$$

and

$$a_n\, f_n(a_n) = b_n\, f_n(b_n). \tag{2.49}$$

Building upon them in same way as for minimum length intervals we arrive to a slightly different set of equations

$$\frac{d\,\phi_1(t)}{dt}\, f_{n+4}\left\{ \frac{1}{\phi_1(t)} \right\} F_p\left\{ \frac{t}{\phi_1(t)} \right\} = \frac{d\,\phi_2(t)}{dt}\, f_{n+4}\left\{ \frac{1}{\phi_2(t)} \right\} F_p\left\{ \frac{t}{\phi_2(t)} \right\} \tag{2.50}$$

$$f_{n+2}\left\{ \frac{1}{\phi_1(t)} \right\} F_p\left\{ \frac{\tau(t)}{\phi_1(t)} \right\} = f_{n+2}\left\{ \frac{1}{\phi_2(t)} \right\} F_p\left\{ \frac{\tau(t)}{\phi_2(t)} \right\} \tag{2.51}$$

with initial conditions

$$\lim_{t\to\infty} \phi_1(t) = \frac{1}{b_n} \text{ and } \lim_{t\to\infty} \phi_2(t) = \frac{1}{a_n}. \tag{2.52}$$

The ratio of the endpoints can be shown to be smaller than the ratio of Tate and Klett intervals while the coverage probability is maintained above $1-\alpha$.

3. General scale parameter. Now we state the assumptions needed in the general scale parameter case, several of which are similar to Brown's (1968) distributional assumptions.

Let Y, S, Z be random variables taking the values y, s, z in $\Re^1 \times \Re^1 \times \Re^q$, $q \geq 0$ and $s > 0$. The variable Z is an ancillary statistic and may not exist. If this is the case, we take $q = 0$. We assume that the distribution of Z, f(z), does not depend on any unknown parameters. Given $Z = z$, the random variables Y and S have a conditional density with respect to Lebesgue measure of the form

$$f_{\mu,\sigma}(s, y \mid z) = \frac{1}{\sigma^2}\, f_{0,1}\left(\frac{s}{\sigma}, \frac{y-\mu}{\sigma} \mid z\right) \tag{3.1}$$

that is, the conditional density belongs to the location-scale family. The location parameter is denoted by μ and the scale parameter by σ. The random variable S has a density

$$f_\sigma(s \mid z) = \frac{1}{\sigma}\, f\left(\frac{s}{\sigma} \mid z\right), \tag{3.2}$$

for some functions f_σ and f, where f is independent of μ.

We are interested in estimating the parameter σ^ρ where $\rho > 0$ a fixed known constant. Any location and scale invariant estimator is of the form $\psi(z)\,s^\rho$, for some function ψ. However, if we require invariance only under rescaling and change of sign of the data, the class of estimators is increased to those of the form $\phi(\,|\,y\,|/s\,,\,z)s^\rho$, for some function ϕ.

When constructing confidence intervals, we consider only connected confidence intervals. Although our results extend to confidence sets that are not connected, such procedures are intuitively unappealing. A sufficient, but not necessary, condition for the minimum length intervals to be connected is that the densities be unimodal, however, we will not make such an assumption. If there is no unimodality, it should be understood that by "minimum length intervals" we mean "minimum length connected intervals".

We need some additional assumptions about the density $f_\sigma(s\,|\,z)$ in order to determine the endpoints of the shortest location and scale invariant $1-\alpha$ confidence intervals for s^ρ. The confidence interval is invariant if it is of the form $(\,\psi_1(z)s^\rho,\ \psi_2(z)s^\rho\,)$, and to have the confidence coefficient to be $1-\alpha$, we must have

$$P\{\,\sigma^\rho \in (\,\psi_1(z)s^\rho,\ \psi_2(z)s^\rho\,)\,\} \geq 1-\alpha. \tag{3.3}$$

Conditioning on z the above inequality becomes

$$\int P\{\,\psi_1(z)s^\rho \leq \sigma^\rho \leq \psi_2(z)s^\rho\,)\,|\,z\,\}\,f(z)\,dz \geq 1-\alpha. \tag{3.4}$$

Observe that the conditional probability does not depend on any unknown parameters since f(z) of (3.2) is free of unknown parameters. Straightforward calculation shows that a sufficient condition for (3.2) is

$$\int_{\frac{1}{\psi_2(z)}}^{\frac{1}{\psi_1(z)}} f(s^\rho\,|\,z)\,ds^\rho = 1-\alpha \tag{3.5}$$

for almost all z. Although equation (3.5) is not necessary for (3.3), if the endpoints $\psi_1(z)$ and $\psi_2(z)$ satisfy (3.3) but not (3.5), then the interval would have undesirable conditional properties since the conditional coverage probability would be bounded on one side of $1-\alpha$ for a range of z values.

Now we derive a convenient expression for the endpoints of the shortest location-scale invariant interval satisfying the probability constraint.

THEOREM 3.1. If $f_\sigma(s \mid z)$ is continuous and has connected support for almost all z then the minimum length location-scale invariant interval, subject to the condition (3.5), is given by

$$C_U(s, z) = (\psi_1(z)s^\rho, \psi_2(z)s^\rho) \tag{3.6}$$

where $\psi_1(z)$ and $\psi_2(z)$ satisfy equation (3.5) and

$$\left\{\frac{1}{\psi_1(z)}\right\}^2 f\left\{\frac{1}{\psi_1(z)} \mid z\right\} = \left\{\frac{1}{\psi_2(z)}\right\}^2 f\left\{\frac{1}{\psi_2(z)} \mid z\right\}. \tag{3.7}$$

PROOF. Equation (3.7) can be derived by using the technique of Lagrange multipliers, minimizing $\psi_2(z) - \psi_1(z)$ subject to the constraint (3.5). In order to use the Lagrange multipliers we need the differentiability of the upper limit of integration as a function of the lower limit, where the integrand is the density of s^ρ. If the assumption of continuity and connected support holds for $f_\sigma(s \mid z)$ then it also holds for $f(s^\rho \mid z)$ and we apply Lemma A.6 to derive equation (3.7). Comparing the difference $\psi_2(z) - \psi_1(z)$ to the possibly multiple solutions of (3.5) and (3.7) guarantees that we have a minimum. □

The assumptions of continuity and connectedness of support of $f_\sigma(s \mid z)$ may be stronger than necessary, because the shortest $1-\alpha$ interval may exist even if they are not met. However, we do not consider it as a major drawback since most densities of practical interest satisfy these conditions.

Now we look for intervals that are superior to interval $C_U(s, z)$ by no longer requiring the interval to be location and scale invariant. The manner of improving is similar to that used in Section 2, shifting the endpoints of the interval whenever $|y|/s$ seems small. In order to have consistent notation we define the statistic $t = y^2/s^2$. Observe that there is a one-to-one relation between t and $|y|/s$.

In the proof of Theorem 3.2 we will assume that given z, y^2 and s are conditionally independent. Even though this does not seem to be a crucial condition for the construction to work, if it holds, the distribution of y^2/σ^2 depends on the parameters only through μ^2/σ^2. Therefore by taking, without loss of generality, $\sigma = 1$ the distribution depends only on $|\mu|$. We will denote the cumulative distribution function of y^2 by $F_\mu(y \mid z)$.

Two additional assumptions, which are important for the construction, are that the distribution $F_\mu(y \mid z)$ has the monotone likelihood ratio property in μ and that it is continuous as a function of y.

For fixed constant K define the interval $I_1(s,t,z,K)$ as follows:

$$I_1(s,t,z,K) = \begin{cases} (\psi_1(z)s^\rho,\ \psi_2(z)s^\rho) & \text{if} \quad t > K \\ (\phi_1(K,z)s^\rho,\ \phi_2(K,z)s^\rho) & \text{if} \quad t \le K \end{cases}$$

where $\phi_1(K,z)$ and $\phi_2(K,z)$ satisfy

$$\int_{\phi_1(K,z)}^{\phi_2(K,z)} \left(\tfrac{1}{x}\right)^2 f\left(\tfrac{1}{x} \middle| z\right) F_0\left(\tfrac{K}{x^{2/\rho}} \middle| z\right) dx = \int_{\psi_1(z)}^{\psi_2(z)} \left(\tfrac{1}{x}\right)^2 f\left(\tfrac{1}{x} \middle| z\right) F_0\left(\tfrac{K}{x^{2/\rho}} \middle| z\right) dx \tag{3.8}$$

and $\phi_1(K,z) \le \psi_1(z)$. Then we have the following theorem.

THEOREM 3.2. The coverage probability of $I_1(s,t,z,K)$ is no less than the coverage probability of $C_U(s,z) = (\psi_1(z)s^\rho, \psi_2(z)s^\rho)$.

PROOF. If $P_{\mu,1}\{t \le K\} = 0$ there is nothing to prove. Otherwise as in Theorem 2.1 we will work with the joint probability

$$P\{\sigma^\rho \in I_1(s,t,z,K),\ t \le K\} = P\left\{\frac{1}{\phi_1(K,z)} \le \frac{s^\rho}{\sigma^\rho} \le \frac{1}{\phi_2(K,z)},\ \frac{y^2}{s^2} \le K\right\}. \tag{3.9}$$

Taking $\sigma = 1$ without loss of generality and conditioning on $s^\rho = w$, yields

$$P\{\sigma^\rho \in I_1(s,t,z,K),\ t \le K\} = \int_{\frac{1}{\phi_2(K,z)}}^{\frac{1}{\phi_1(K,z)}} f(w\,|\,z)\ P_\mu(y^2 \le Kw^{2/\rho}\,|\,s^\rho = w,\ z)\ dw$$

$$= \int_{\frac{1}{\phi_2(K,z)}}^{\frac{1}{\phi_1(K,z)}} f(w\,|\,z)\ F_\mu(\,Kw^{2/\rho}\,|\,z\,)\ dw. \tag{3.10}$$

Using the transformation $x = 1/w$ and the observation that the intervals $I_1(s,t,z,K)$ and $C_U(s,z)$ differ only when $t \le K$, we see that the theorem is proved if we establish

$$\int_{\phi_1(K,z)}^{\phi_2(K,z)} \left(\tfrac{1}{x}\right)^2 f(\tfrac{1}{x}\,|\,z)\ F_\mu(\,\tfrac{K}{x^{2/\rho}}\,|\,z\,)\ dx \ge \int_{\psi_1(z)}^{\psi_2(z)} \left(\tfrac{1}{x}\right)^2 f(\tfrac{1}{x}\,|\,z)\ F_\mu(\,\tfrac{K}{x^{2/\rho}}\,|\,z\,)\ dx \tag{3.11}$$

for every μ. The rest of the proof is similar to the proof of the Theorem 2.1. In order to differentiate the upper limit of integration as a function of the lower limit we need the integrand $(1/x)^2 f(1/x\,|\,z)\ F_0(K/x^{2/\rho}\,|\,z)$ to be continuous, which follows from the assumptions of the continuity of $f(1/x\,|\,z)$ and $F_0(K/x^{2/\rho}\,|\,z)$. Furthermore, since $F_0(K/x^{2/\rho}\,|\,z)$ is monotone and $f(1/x\,|\,z)$ has connected support, the integrand has connected support. Hence we can apply Lemma A.6 and derive an expression for the derivative.

The other key assumption needed is the monotone likelihood ratio of F_μ. After obtaining the derivative, we are lead to an expression analogous to (2.15) and the monotone likelihood ratio property of F_μ completes the proof. □

REMARK. If y^2 and s are not conditionally independent, the probability $P_\mu(y^2 \le Kw^{2/\rho}\,|\,s^\rho = w,\ z)$ may not have a tractable form. However, making the conditional independence assumption we have

$$P_\mu(y^2 \leq Kw^{2/\rho} \mid s^\rho = w, z) = F_\mu(Kw^{2/\rho} \mid z) \tag{3.12}$$

which justifies (3.10). The assumption is used only to ensure that the argument of F_μ is a monotone function of w. Our construction would work if, instead of conditional independence and monotone likelihood ratio of F_μ, we require the function

$$\frac{P_\mu(y^2 \leq Kw^{2/\rho} \mid s^\rho = w, z)}{P_0(y^2 \leq Kw^{2/\rho} \mid s^\rho = w, z)} \tag{3.13}$$

to be increasing in w.

We saw that for every $\phi_1(K,z)$ and $\phi_2(K,z)$ satisfying (3.8) and $\phi_1(K,z) \leq \psi_1(z)$ the coverage probability of $I_1(s,t,z,K)$ is at least $1-\alpha$. However, in order to gain in length we need some additional restrictions on $\phi_1(K,z)$. When $t \leq K$, the length of the interval is equal to $(\phi_2(K,z) - \phi_1(K,z))s^\rho$. Subject to (3.8) the length is decreasing as a function of the lower limit of integration if $dg_{\gamma_1,0}(w)/dw > 1$. Using the formula for the derivative of $g_{\gamma,0}(w)$ derived from Lemma A.6, we obtain the expression

$$\left.\frac{dg_{\gamma,0}(w)}{dw}\right|_{w=\psi_1(z)} = \frac{F_0\left(\frac{K}{\{\psi_1(z)\}^{2/\rho}} \mid z\right)}{F_0\left(\frac{K}{\{\psi_2(z)\}^{2/\rho}} \mid z\right)}. \tag{3.14}$$

The last ratio is always greater than or equal to one since $\psi_1(z) < \psi_2(z)$ and F_0 is a nondecreasing function. Therefore we cannot have an increase on the length for any K. Lemma A.7 shows that we can always chose an appropriate K such that

$$F_0\left(\frac{K}{\{\psi_1(z)\}^{2/\rho}} \mid z\right) > F_0\left(\frac{K}{\{\psi_2(z)\}^{2/\rho}} \mid z\right) \tag{3.15}$$

that is, the derivative of the length, as a function of the lower limit, is strictly negative in some neighborhood of $\psi_1(z)$. Since from Lemma A.6 we also know that the derivative is continuous, for every $\phi_1(K,z)$ sufficiently close to $\psi_1(z)$ we will have some gain in length. We now make more precise what we mean by "sufficiently close".

Define $\phi_1^0(K,z)$ and $\phi_2^0(K,z)$ to satisfy

$$\int_{\phi_1^0(K,z)}^{\phi_2^0(K,z)} \left(\frac{1}{x}\right)^2 f\left(\frac{1}{x} \,|\, z\right) F_0\left(\frac{K}{x^{2/\rho}} \,|\, z\right) dx = \int_{\psi_1(z)}^{\psi_2(z)} \left(\frac{1}{x}\right)^2 f\left(\frac{1}{x} \,|\, z\right) F_0\left(\frac{K}{x^{2/\rho}} \,|\, z\right) dx \tag{3.16}$$

and

$$\left(\frac{1}{\phi_1^0(K,z)}\right)^2 f\left(\frac{1}{\phi_1^0(K,z)} \,|\, z\right) F_0\left(\frac{K}{\{\phi_1^0(K,z)\}^{2/\rho}} \,|\, z\right)$$
$$= \left(\frac{1}{\phi_2^0(K,z)}\right)^2 f\left(\frac{1}{\phi_2^0(K,z)} \,|\, z\right) F_0\left(\frac{K}{\{\phi_2^0(K,z)\}^{2/\rho}} \,|\, z\right). \tag{3.17}$$

Since we do not assume unimodality of the function $(1/x)^2 f(1/x \mid z)\ F_0(K/x^{2/\rho} \mid z)$ there may be more than one solution to (3.16) and (3.17). If there are no solutions with $\phi_1^0(K,z) \leq \psi_1(z)$, then any $\phi_1(K,z) \leq \psi_1(z)$ and $\phi_2(K,z)$ satisfying (3.8) defines a confidence interval $I_1(s,t,z,K)$ shorter than $(\psi_1(z)s^\rho,\ \psi_1(z)s^\rho)$. Otherwise we take $\phi_1(K,z)$ greater than every solution to equations (3.16) and (3.17) with $\phi_1^0(K,z) \leq \psi_1(z)$. Then, defining

$$\gamma = \int_{\psi_1(z)}^{\psi_2(z)} \left(\frac{1}{x}\right)^2 f\left(\frac{1}{x} \,|\, z\right) F_0\left(\frac{K}{x^{2/\rho}} \,|\, z\right) dx, \tag{3.18}$$

we have $dg_{\gamma,0}(w)/dw > 1$ for every w in the interval $(\ \phi_1(K,z),\ \psi_1(z)\ )$ which implies that we have some gain in terms of length. Thus we need to take $\phi_1(K,z)$ close enough to $\psi_1(z)$ so

$$\left(\frac{1}{w}\right)^2 f\left(\frac{1}{w} \,|\, z\right) F_0\left(\frac{K}{w^{2/\rho}} \,|\, z\right) > \left(\frac{1}{g_{\gamma,0}(w)}\right)^2 f\left(\frac{1}{g_{\gamma,0}(w)} \,|\, z\right) F_0\left(\frac{K}{\{g_{\gamma,0}(w)\}^{2/\rho}} \,|\, z\right)$$
$$\forall\ w \in [\ \phi_1(K,z),\ \psi_1(z)\ ], \tag{3.19}$$

for γ in (3.18). The last requirement is intuitively expected. If the value of the integrand at the lower limit of the interval is smaller that the value at the upper limit then by keeping the area constant and shifting the lower limit towards zero we would increase the distance between the endpoints.

If $\phi_1(K,z)$ is close enough to $\psi_1(z)$ to satisfy both (3.8) and (3.19), we can further improve upon the interval $I_1(s,t,z,K)$ by taking another cutoff point. By further partitioning the set of possible values of t, we can construct confidence intervals that are based on more cutoff points, and eventually fill the interval $(0, +\infty)$ with points. The intervals based on $\underset{\sim}{K}_1, \underset{\sim}{K}_2, \ldots$ will be denoted by $I_m(s,t,z,\underset{\sim}{K}_m)$ and the endpoints satisfy:

For $i = 1, 2, \ldots, m-1$,

$$\int_{\phi_1(K_{m,i},z)}^{\phi_2(K_{m,i},z)} \left(\frac{1}{x}\right)^2 f\left(\frac{1}{x} \mid z\right) F_0\left(\frac{K_{m,i}}{x^{2/\rho}} \mid z\right) dx$$

$$= \int_{\phi_1(K_{m,i+1},z)}^{\phi_2(K_{m,i+1},z)} \left(\frac{1}{x}\right)^2 f\left(\frac{1}{x} \mid z\right) F_0\left(\frac{K_{m,i}}{x^{2/\rho}} \mid z\right) dx. \tag{3.20}$$

For $i = m$,

$$\int_{\phi_1(K_{m,m},z)}^{\phi_2(K_{m,m},z)} \left(\frac{1}{x}\right)^2 f\left(\frac{1}{x} \mid z\right) F_0\left(\frac{K_{m,m}}{x^{2/\rho}} \mid z\right) dx$$

$$= \int_{\psi_1(z)}^{\psi_2(z)} \left(\frac{1}{x}\right)^2 f\left(\frac{1}{x} \mid z\right) F_0\left(\frac{K_{m,m}}{x^{2/\rho}} \mid z\right) dx, \tag{3.21}$$

where $\phi_1(K_{m,1},z) \leq \phi_1(K_{m,2},z) \leq \cdots \leq \phi_1(K_{m,m},z) \leq \psi_1(z)$ and also for each $i = 1, \ldots, m$

$$\left(\frac{1}{w}\right)^2 f\left(\frac{1}{w} \mid z\right) F_0\left(\frac{K_{m,i}}{w^{2/\rho}} \mid z\right) > \left(\frac{1}{g_{\gamma_i,0}(w)}\right)^2 f\left(\frac{1}{g_{\gamma_i,0}(w)} \mid z\right) F_0\left(\frac{K_{m,i}}{\{g_{\gamma_i,0}(w)\}^{2/\rho}} \mid z\right)$$

$$\forall\ w \in [\,\phi_1(K_{m,i},z),\ \phi_1(K_{m,i+1},z)\,] \tag{3.22}$$

where

$$\gamma_i = \int_{\phi_1(K_{m,i+1},z)}^{\phi_2(K_{m,i+1},z)} \left(\frac{1}{x}\right)^2 f\left(\frac{1}{x} \mid z\right) F_0\left(\frac{K_{m,i}}{x^{2/\rho}} \mid z\right) dx. \tag{3.23}$$

If there is no $w \in [\,\phi_1(K_{m,i},z),\ \phi_1(K_{m,i+1},z)\,]$ such that (3.22) is satisfied then we take $\phi_1(K_{m,i},z) = \phi_1(K_{m,i+1},z)$.

The intervals $I_m(s,t,z,\underset{\sim}{K}_m)$ have minimum coverage probability equal to $1-\alpha$. It is not guaranteed that for a given m, the intervals are shorter than $C_U(s, z)$ since we may have

$$P\{ t \le K_{m,m} \} = 0, \tag{3.24}$$

or

$$F_0\Big(\frac{K_{m,i}}{\{\psi_1(z)\}^{2/\rho}} \mid z \Big) = F_0\Big(\frac{K_{m,i}}{\{\psi_2(z)\}^{2/\rho}} \mid z \Big) \tag{3.25}$$

for $i = 1, 2, \ldots, m$. In the first case the intervals $C_U(s, z)$ and $I_m(s,t,z, \underset{\sim}{K}_m)$ coincide, whereas in the second case the intervals have different endpoints but the lengths are the same. However, equation (3.24) cannot hold for every m, because it would mean that $P\{ t < +\infty \} = 0$, since $\lim_{m\to\infty} K_{m,m} = +\infty$. On the other hand Lemma A.7 shows that we cannot have

$$F_0\Big(\frac{K}{\{\psi_1(z)\}^{2/\rho}} \mid z \Big) = F_0\Big(\frac{K}{\{\psi_2(z)\}^{2/\rho}} \mid z \Big) \tag{3.26}$$

for every K. Therefore, by filling $(0, +\infty)$ with cutoff points, we know that, eventually, the interval $I_m(s,t,z, \underset{\sim}{K}_m)$ will improve upon $C_U(s, z)$.

As $m \to \infty$ the endpoints of $I_m(s,t,z, \underset{\sim}{K}_m)$ tend to some functions $\phi_1(t,z)$ and $\phi_2(t,z)$. In order to determine $\phi_1(t,z)$ and $\phi_2(t,z)$ we work as in Section 2, assuming that the function $F_0(x \mid z)$ is twice differentiable with second derivative bounded in finite intervals. That is the only additional assumption we need to derive the equation

$$\frac{d\phi_1(t,z)}{dt} \Big(\frac{1}{\phi_1(t,z)} \Big)^2 f\Big(\frac{1}{\phi_1(t,z)} \mid z \Big) F_0\Big(\frac{t}{\{\phi_1(t,z)\}^{2/\rho}} \mid z \Big)$$
$$= \frac{d\phi_2(t,z)}{dt} \Big(\frac{1}{\phi_2(t,z)} \Big)^2 f\Big(\frac{1}{\phi_2(t,z)} \mid z \Big) F_0\Big(\frac{t}{\{\phi_2(t,z)\}^{2/\rho}} \mid z \Big). \tag{3.27}$$

Note that the last relation does not determine $\phi_1(t,z)$ and $\phi_2(t,z)$. In general we cannot specify another equation that uniquely defines the endpoints because requirement (3.22) does not uniquely determine $I_m(s,t,z, \underset{\sim}{K}_m)$.

By the Lebesgue Dominated Convergence Theorem we know that the confidence coefficient of any interval constructed in this way is $1-\alpha$. We also saw that for every finite step m, possibly for m greater than some m_0, the interval $\mathbf{I_m}(s,t,z,\underset{\sim}{K}_m)$ is shorter than $C_U(s, z)$. However it is not clear what happens with the limiting interval. Only in special cases we can specify another equation that defines the limiting interval and make statements about its length.

4. Numerical results. We now investigate numerically the gains in coverage probability and expected length of the confidence intervals for the variance of normal, constructed in Section 2. Previous relative risk calculations for the point estimator (Rukhin 1987) and the numerical results of Shorrock (1987) suggest that the improvement can only be minimal for small values of p.

The gains are substantial for small and moderate n and for p large relatively to n. The endpoints of the intervals depend on a rather arbitrary function $\tau(t)$, and the numerical results show that dependence of both coverage probability and expected length on $\tau(t)$ is rather strong. For different functional forms we have little feeling about what to expect, so we chose a wide variety of $\tau(t)$ forms. The functional forms of $\tau(t)$ which seem to be optimal are moderately or slowly increasing. Rapidly increasing $\tau(t)$ have an effect only when n is small. For large n only moderately increasing $\tau(t)$ can change the coverage probability and the length substantially.

In figure 1 we see that the largest relative gain in length that we obtained was about 5.3%. Figure 2 shows that the largest difference between coverage probability and confidence coefficient was about .0033. The wide selection of $\tau(t)$ makes it difficult to find an optimal functional form and suggests that there may be other forms that perform better. However, among the intervals we computed, none dominates the others in both coverage probability and length.

The coverage probability and expected length were calculated by numerical integration. The computations were performed on Purdue University Computing Center's IBM 3090-180E Computer using FORTRAN programming language and IMSL subroutines. The graphs were produced on Cornell University's IBM 3090-200 computer using SAS/GRAPH.

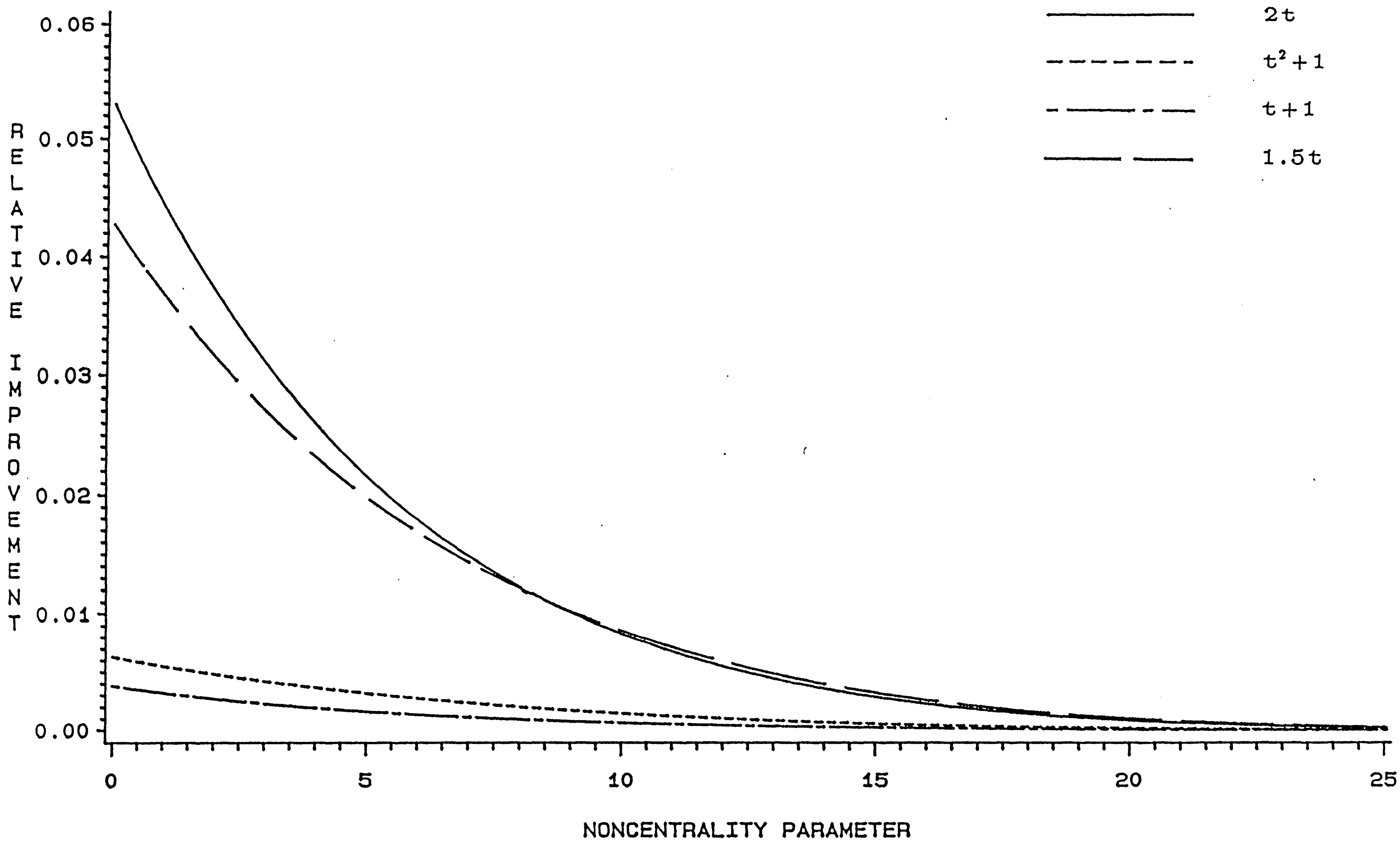

FIGURE 1: EXPECTED RELATIVE IMPROVEMENT IN LENGTH FOR N=25 AND P=10
THE CONFIDENCE COEFFICIENT IS 0.95

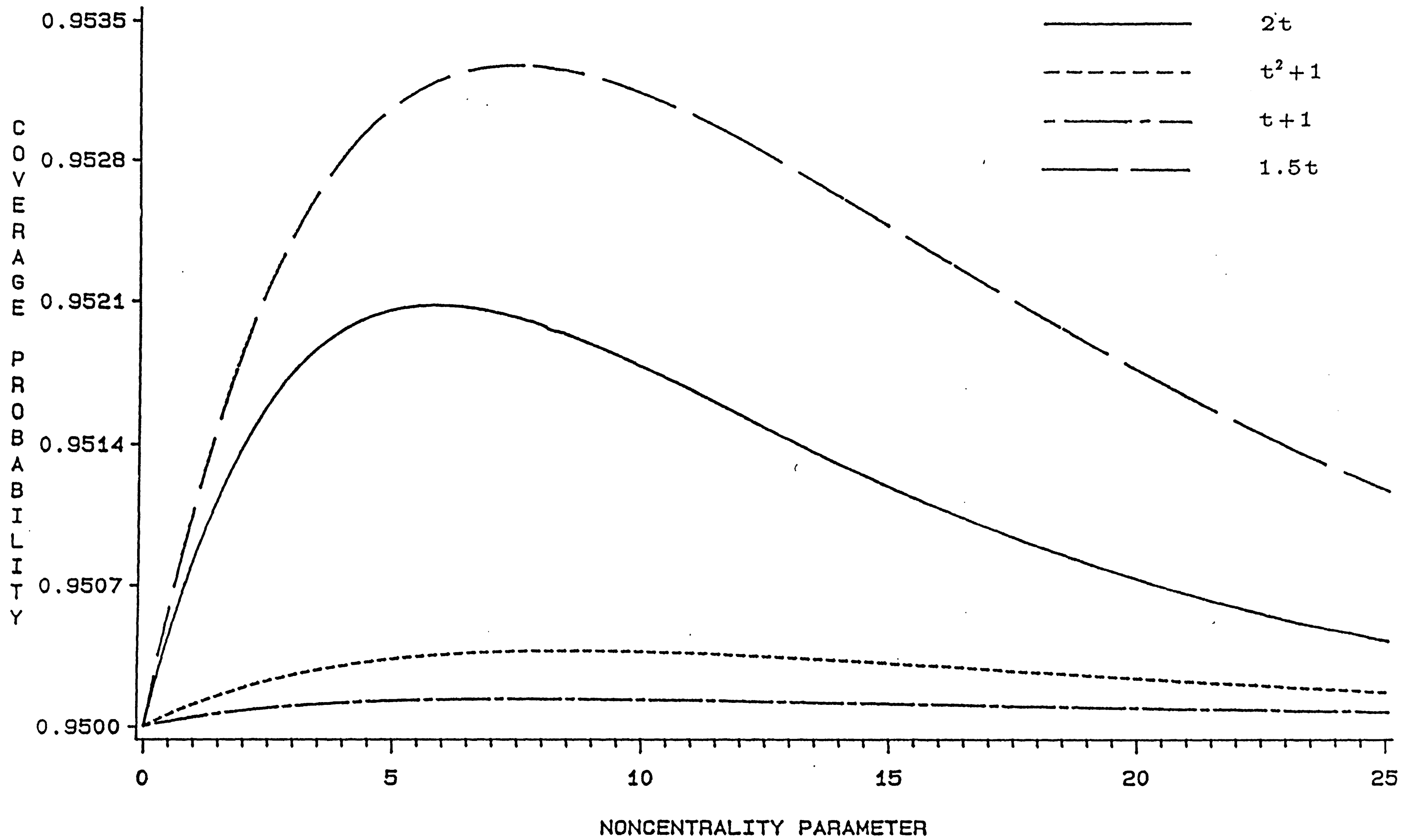

FIGURE 2: COVERAGE PROBABILITY FOR N=25 AND P=10
THE CONFIDENCE COEFFICIENT IS 0.95

APPENDIX

LEMMA A.1. If a differentiable function $f(x)$ defined on the real line has $f'(x) < 0$ whenever $f(x) = 0$ and there is an x_0 such that $f(x_0) = 0$, then $f(x)$ is positive for $x < x_0$ and negative for $x > x_0$.

LEMMA A.2. If $\phi_1(K)$ is defined (2.6) and (2.7) then $\phi_1(K) < 1/b_n$ for any choice of τ and K.

PROOF. From equation (2.6) we see that neither of the sets (ϕ_1, ϕ_2) and $(1/b_n, 1/a_n)$ can be a proper subset of the other. Therefore we may have the following cases:

(i) $\phi_1 < \frac{1}{b_n} < \phi_2 < \frac{1}{a_n}$

(ii) $\phi_1 < \phi_2 \leq \frac{1}{b_n} < \frac{1}{a_n}$

(iii) $\frac{1}{b_n} < \phi_1 < \frac{1}{a_n} < \phi_2$

(iv) $\frac{1}{b_n} < \frac{1}{a_n} \leq \phi_1 < \phi_2$

(v) $\frac{1}{b_n} = \phi_1 < \frac{1}{a_n} = \phi_2$

We will show that cases (iii), (iv) and (v) are vacuous. The unimodality of $f_{n+4}(1/x)\, F_p(\tau(K)/x)$ as a function of x, by Lemma A.4, and equation (2.7) imply that for every $x \in (\phi_1, \phi_2)$

$$f_{n+4}\left(\frac{1}{x}\right) F_p\left(\frac{\tau(K)}{x}\right) > f_{n+4}\left(\frac{1}{\phi_1}\right) F_p\left(\frac{\tau(K)}{\phi_1}\right) \tag{A.1}$$

and for every $x \in (0, \phi_1)$

$$f_{n+4}\left(\frac{1}{x}\right) F_p\left(\frac{\tau(K)}{x}\right) < f_{n+4}\left(\frac{1}{\phi_1}\right) F_p\left(\frac{\tau(K)}{\phi_1}\right) \tag{A.2}$$

and $f_{n+4}(1/x)\, F_p(\tau(K)/x)$ is increasing in x.

If $1/b_n < \phi_1 < 1/a_n$ then equations (A.1) and (A.2) imply

$$f_{n+4}(b_n)\, F_p(\tau(K) b_n) < f_{n+4}(a_n)\, F_p(\tau(K) a_n) \tag{A.3}$$

which contradicts $f_{n+4}(b_n) = f_{n+4}(a_n)$ since $F_p(\tau(K)x)$ is increasing in x and $b_n > a_n$. Hence case (iii) is not possible.

If $1/b_n < 1/a_n \leq \phi_1$, we can conclude (A.3) because $f_{n+4}(1/x)\, F_p(\tau(K)/x)$ is increasing in x for $x < \phi_1$, so case (iv) is not possible either.

If $1/b_n = \phi_1 < 1/a_n = \phi_2$ then equation (2.7) implies

$$f_{n+4}(b_n)\, F_p(\tau(K)\, b_n) = f_{n+4}(a_n)\, F_p(\tau(K)\, a_n) \tag{A.4}$$

which also contradicts $f_{n+4}(b_n) = f_{n+4}(a_n)$. Therefore the only possible cases are (i) and (ii), that is, we must have $\phi_1(K) < 1/b_n$. □

LEMMA A.3. If $\phi_1(K_1)$ and $\phi_1(K_2)$ are defined by equations (2.21) – (2.23), then $\phi_1(K_1) > \phi_1(K_2)$.

PROOF. The proof is similar to the proof of Lemma A.2. There are five possible cases and using the unimodality of $f_{n+4}(1/x)\, F_p(\tau(K_2)/x)$, by Lemma A.4, and equation (2.23) with $i = 2$ we can see that for $x \in (\, \phi_1(K_2),\, \phi_2(K_2)\,)$

$$f_{n+4}\left\{ \frac{1}{x} \right\} F_p\left\{ \frac{\tau(K_2)}{x} \right\} > f_{n+4}\left\{ \frac{1}{\phi_1(K_1)} \right\} F_p\left\{ \frac{\tau(K_2)}{\phi_1(K_1)} \right\} \tag{A.5}$$

whereas for $x \in (\, 0,\, \phi_1(K_2)\,)$

$$f_{n+4}\left\{ \frac{1}{x} \right\} F_p\left\{ \frac{\tau(K_2)}{x} \right\} < f_{n+4}\left\{ \frac{1}{\phi_1(K_1)} \right\} F_p\left\{ \frac{\tau(K_2)}{\phi_1(K_1)} \right\} \tag{A.6}$$

and $f_{n+4}(1/x)\, F_p(\tau(K_2)/x)$ is increasing in x. Cases (iii), (iv) and (v) of Lemma A.2 imply, in a similar way, that

$$f_{n+4}\left\{ \frac{1}{\phi_1(K_1)} \right\} F_p\left\{ \frac{\tau(K_2)}{\phi_1(K_1)} \right\} \geq f_{n+4}\left\{ \frac{1}{\phi_2(K_1)} \right\} F_p\left\{ \frac{\tau(K_2)}{\phi_2(K_1)} \right\} \tag{A.7}$$

which is analogous to equation (A.3). Now if we use (2.23) with $i = 1$ we see that (A.7) is equivalent to

$$\frac{F_p\left\{ \frac{\tau(K_2)}{\phi_1(K_1)} \right\}}{F_p\left\{ \frac{\tau(K_1)}{\phi_1(K_1)} \right\}} \leq \frac{F_p\left\{ \frac{\tau(K_2)}{\phi_2(K_1)} \right\}}{F_p\left\{ \frac{\tau(K_1)}{\phi_2(K_1)} \right\}}. \tag{A.8}$$

Applying Lemma A.5 with $x_1 = \tau(K_1)/\phi_1(K_1)$, $x_2 = \tau(K_1)/\phi_2(K_1)$ and $\beta = \tau(K_2)/\tau(K_1)$

(which is smaller than 1 since the function τ is increasing) we get

$$\frac{F_p\left\{\frac{\tau(K_2)}{\phi_1(K_1)}\right\}}{F_p\left\{\frac{\tau(K_1)}{\phi_1(K_1)}\right\}} > \frac{F_p\left\{\frac{\tau(K_2)}{\phi_2(K_1)}\right\}}{F_p\left\{\frac{\tau(K_1)}{\phi_2(K_1)}\right\}} \tag{A.9}$$

which is the necessary contradiction. □

LEMMA A.4. If f_k and F_k denote respectively the chi squared probability density and cumulative distribution function with k degrees of freedom, then for any integers n and p and positive constant M the function $f_{n+4}(1/x)\, F_p(M/x)$ is a unimodal function of x.

PROOF. It follows from the fact that unimodality is preserved under monotone transformation of the argument and Lemma A.4 of Shorrock (1987), which states that for $n \geq 2$ and $p \geq 1$ the function $f_n(x)\, F_p(Mx)$ is unimodal. □

LEMMA A.5. Let F_p be a chi squared distribution function with $p \geq 1$ degrees of freedom. If $\beta < 1$ and $x_1 > x_2$, then

$$\frac{F_p(\beta x_1)}{F_p(x_1)} > \frac{F_p(\beta x_2)}{F_p(x_2)}. \tag{A.10}$$

PROOF. It follows from the fact that the gamma densities have monotone likelihood ratio in the scale parameter. (See also Lemma 4.2 of Cohen (1972)). Note that property (A.10) implies that the distribution is stochastically increasing in the parameter. □

LEMMA A.6. Let f(x) be a nonnegative integrable function on the real line and γ a fixed constant smaller than $\int_{-\infty}^{\infty} f(x)\, dx$. Let

$$D = \{\, w \mid \int_{w}^{\infty} f(x)\, dx > \gamma \,\} \tag{A.11}$$

and $g_\gamma(w) : D \rightarrow \Re$ be defined as the solution to the equation

$$\gamma = \int_{w}^{g_\gamma(w)} f(x)\, dx. \tag{A.12}$$

If f(x) is continuous and the set $E = \{\, x \mid f(x) \neq 0 \,\}$ is connected then the function g_γ is differentiable and its derivative is equal to $f(w)/f\{g_\gamma(w)\}$. Furthermore, the derivative is continuous.

PROOF. Observe that the function $g_\gamma(w)$ is continuous. If it were not, since it is increasing, it must have a jump, that is

$$U = \lim_{w \downarrow w_0} g_\gamma(w) > \lim_{w \uparrow w_0} g_\gamma(w) = L \tag{A.13}$$

which implies that

$$\int_L^U f(x)\, dx = 0. \tag{A.14}$$

But the last relation contradicts the assumption that E is connected.

Fix $w_0 \in D$ and let $w < w_0$ Then

$$\int_{w_0}^{g_\gamma(w_0)} f(x)\, dx = \int_{w}^{g_\gamma(w)} f(x)\, dx = \gamma \tag{A.15}$$

implies

$$\int_{w}^{w_0} f(x)\, dx = \int_{g_\gamma(w)}^{g_\gamma(w_0)} f(x)\, dx. \tag{A.16}$$

If $M_1 = \sup \{ f(x) \mid w \le x \le w_0 \}$ and

$N_1 = \inf \{ f(x) \mid g_\gamma(w) \le x \le g_\gamma(w_0) \}$ then we have

$$\int_{w}^{w_0} f(x)\, dx \le M_1 (w_0 - w) \tag{A.17}$$

and

$$\int_{g_\gamma(w)}^{g_\gamma(w_0)} f(x)\, dx \ge N_1 \, \{g_\gamma(w_0) - g_\gamma(w)\}. \tag{A.18}$$

The last two equations, together with (A.16), imply that

$$\frac{M_1}{N_1} \ge \frac{g_\gamma(w_0) - g_\gamma(w)}{w_0 - w}. \tag{A.19}$$

Note that N_1 is strictly positive because it is the infimum of the function f(x) over the closed interval [$g_\gamma(w)$, $g_\gamma(w_0)$] and f(x) is strictly positive for every $x \in [\, g_\gamma(w), g_\gamma(w_0)\,]$. If $f(x_0) = 0$ for some x_0, then, since the set E is connected, it would be zero for every $x > x_0$. That would imply that

$$\int_{x_0}^{\infty} f(x)\, dx = \int_{g_\gamma(w_0)}^{\infty} f(x)\, dx = 0 \tag{A.20}$$

which contradicts $w_0 \in D$. Hence the LHS of (A.19) is finite.

On the other hand if we define

$M_2 = \inf \{\, f(x) \mid w \leq x \leq w_0 \,\}$ and

$N_2 = \sup \{\, f(x) \mid g_\gamma(w) \leq x \leq g_\gamma(w_0) \,\}$ then

$$N_2 \,\{g_\gamma(w_0) - g_\gamma(w)\} \geq \int_{g_\gamma(w)}^{g_\gamma(w_0)} f(x)\, dx = \int_{w}^{w_0} f(x)\, dx \geq M_2(w_0 - w) \tag{A.21}$$

which implies

$$\frac{M_2}{N_2} \leq \frac{g_\gamma(w_0) - g_\gamma(w)}{w_0 - w}. \tag{A.22}$$

Putting the equations (A.19) and (A.22) together

$$\frac{M_2}{N_2} \leq \frac{g_\gamma(w_0) - g_\gamma(w)}{w_0 - w} \leq \frac{M_1}{N_1}. \tag{A.23}$$

Letting $w \to w_0$, because of the continuity of f(x) and $g_\gamma(w)$, we have

$$\lim_{w \to w_0} M_1 = \lim_{w \to w_0} M_2 = f(w_0) \tag{A.24}$$

and

$$\lim_{w \to w_0} N_1 = \lim_{w \to w_0} N_2 = f\{g_\gamma(w_0)\}. \tag{A.25}$$

From equations (A.23) – (A.25) we conclude that

$$\lim_{w \to w_0} \frac{g_\gamma(w_0) - g_\gamma(w)}{w_0 - w} = \frac{f(w_0)}{f\{g_\gamma(w_0)\}}. \tag{A.26}$$

Repeating the argument for $w > w_0$ we obtain the same limit. Therefore the function $g_\gamma(w)$ is differentiable and its derivative is equal to $f(w)/f\{g_\gamma(w)\}$. The derivative is continuous as a ratio of two continuous functions. □

LEMMA A.7. Let F be a nondegenerate cumulative distribution function such that $F(0) = 0$. Then for every $x_1 < x_2$, there is a K such that $F(Kx_1) < F(Kx_2)$.

PROOF. Suppose that $F(Kx_1) = F(Kx_2)$ for every K. By letting $K = 1$ we have $F(x_1) = F(x_2)$. Since F is nondecreasing, $F(x) = F(x_1) = F(x_2)$ for every $x \in (x_1, x_2)$. Letting $K = x_1/x_2$ we get $F(x_1) = F(Kx_2) = F(Kx_1)$ therefore $F(x) = F(x_1) = F(x_2)$ for every $x \in (x_1^2 x_2, x_2)$. Repeating the same argument for $K = (x_1/x_2)^2, (x_1/x_2)^3 \ldots$ we can see that F must be constant for every $x \in ((x_1/x_2)^2, x_2), ((x_1/x_2)^3, x_2) \ldots$, hence for every $x < x_2$. In a similar way we can show that F must be constant for every $x > x_1$, that is, it is a degenerate distribution function. □

REFERENCES

BREWSTER, J. F. and ZIDEK, J. V. (1974). Improving on equivariant estimators. *Ann. Statist.* 2 21-38.

BROWN, L. D. (1968). Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters. *Ann. Math. Statist.* 39 29-48.

COHEN, A. (1972). Improved confidence intervals for the variance of a normal distribution. *J. Amer. Statist. Assoc.* 67 382-387.

HODGES, J. L. Jr. and LEHMANN E. L. (1951). Some applications of the Cramér-Rao inequality. *Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California, Berkeley. 13-22.

PROSKIN, H. M. (1985). An admissibility theorem with applications to the estimation of the variance of a normal distribution. Unpublished Ph.D. dissertation, Rutgers University, Department of Statistics.

RUKHIN, A. L. (1987). How much better are better estimators of a normal variance. *J. Amer. Statist. Assoc.* 82 925-928.

SHORROCK, G. (1987). Improved confidence intervals for a normal variance. Technical Report, Université du Québec à Montréal.

STEIN, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Ann. Inst. Statist. Math.* 16 155-160.

TATE, R. F. and KLETT, G. W. (1959). Optimal confidence intervals for the variance of a normal distribution. *J. Amer. Statist. Assoc.* 54 674-682.